Impact of Exam Scheduling on Final Exam Grades within the Department of Chemistry at UBC

Kevin Multani[1], Sharon Mui[1]
Department of Statistics[1], University of British Columbia

DEPARTMENT OF STATISTICS

ABSTRACT

As part of UBC's strategic plan, various teaching and learning initiatives have been launched with the aim to improve students' disciplinary learning ability, motivation, attitude, and communication skills. Two professors in the UBC Department of Chemistry, Dr. Stewart and Dr. Dake have anecdotally observed that final exam scheduling has a strong impact on final exam performance. Consequently, they commissioned our group to identify if there was any statistically significant impact on final exam grades, due to scheduling.

Our analysis has shown that there is sufficient evidence that final exam scheduling has an impact on final exam grades. For example, our model has suggests that if a student were given one more day to study for a final exam, then their grade would move up 0.59%, after controlling numerous modes of variability.

INTRODUCTION

Our clients, Dr. Stewart, and Dr. Dake anecdotally saw that students within the Chemistry department would perform non-optimally if they were subjected to certain types of final exam scheduling. For example, if one of their students had two final exams back to back, one of the exams' grades would suffer greatly. Our team has therefore engaged in a project to understand how students' final exam scores are affected by exam scheduling.

Primary Research Question

Is there *any* effect of exam scheduling on the final exam performance in core Chemistry 200 and 300-level courses for students in the Chemistry or Biochemistry Major after controlling for possible confounding, or other attributable, factors? If so, what are the magnitudes of these effects in terms of final exam percentage points.

METHODS

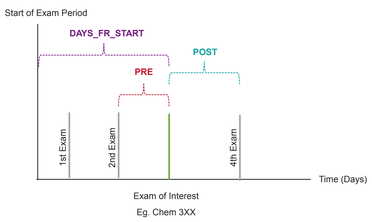
Beyond scheduling, other effects play a role in controlling student final exam performance. Below, we have generalized the primary confounding effects hypothesized by our team.

Exam Scheduling Effects	Description
Scheduling	The impact that final exam scheduling has on a student's final exam performance.
Student	The effect on final exam performance based on differences among students.
Class	The effect on final exam performance based on the difference among different sessions of the same course.
Stress	The effect on final exam performance based on stress-causing factors for students (other courses, personal, work, etc.)

Our clients have provided anonymized data regarding course schedules, professor IDs, and student demographics to model these four scheduling effects. Furthermore, from the available data, our team generated variables (PRE, POST, DAYS_FR_START) as scheduling metrics in our model analysis.

Given a student's grade for a CHEM exam:

- PRE - number of days since last exam
- POST - number of days until next exam
- DAYS_FR_START - days since the beginning of exam period



RESULTS

We used a Mixed Effect Model. The following are the fixed effects:

- Course number (i.e. CHEM 200)
- Session (i.e. 2012W2)
- DAYS_FR_START (number of days since the start of the exam period)
- PRE (number of days available to study)
- POST (number of days until next exam)
- Student's sessional average of the previous year.

The following is the random effect (random intercept):

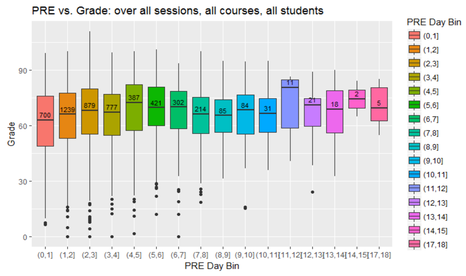
- Student ID

Variable Name	Estimate	P-Value
DAYS_FR_START	0.22	<0.01
PRE	0.59	<0.01
POST	0.20	<0.01
SESS_AVG	0.92	<0.01

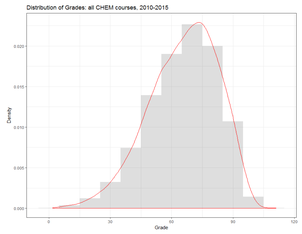
EXPLORATORY

Our exploratory analysis includes:

- Histograms of Student Grades Per Year
- Histogram of all Student Grades
- Plot of Grades vs. Time since last Exam
- Plot of Grades vs. Time until next Exam
- Plot of Grades Across Absolute Time
- Q-Q plot of Grades
- Other plots



The figure above shows boxplots of PRE Day Bin vs. Grade. PRE Day Bin (0,1], for example, means that all observations with a PRE value between 0 and 1 (inclusive) are collected. The solid black line within a boxplot shows the median. The whiskers indicate a 95% confidence interval for the median. The number indicated within a boxplot denotes the number of observations that belong to that PRE Day Bin. (refer to Methods for variable definitions)



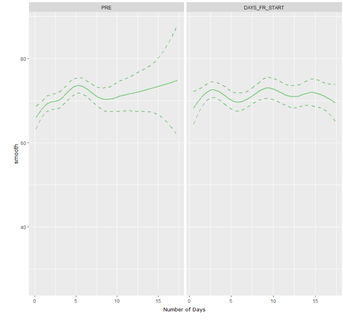
This figure to the left shows the empirical (sample) distribution of the provided chemistry course grades for students within Chemistry and Biochemistry. The *y*-axis is the density, which is the frequency scaled by a constant so that the area of the distribution is equal to 1. Grades refer to the final exam, unscaled grade that a student received. The average grade is 65.7% and the median grade is 67.5%.

General findings from the exploratory analysis:

- We expected a positive, increasing trend in Grade as PRE value increased. Our results showed a non-linear effect, lower variance, and slight increase in grades across PRE.
- To satisfy underlying assumptions in our statistical models, the true distribution of our response variable, grades, should be normally distributed. Our empirical distribution shows a left skew in the data. We've noted these results and have accordingly, approached our results interpretation with caution.

DISCUSSION

This figure shows an initial investigation on the nonlinearity of the data. Both trends are statistically significant ($p < 0.01$ for PRE and $p = 0.01$ for DAYS_FR_START). Note here, the *y*-axis does not represent the raw grade. All we care about is the shape of plots. The dashed lines represent 95% confidence bands for the curve.



There is evidence suggesting that the true model is non-linear (see Figure 7). Further analysis can be focused on the non-linearity of those relationships, e.g., imposing cubic relationship between numerical variables and the response. Investigating nonlinearity would only extend the result, meaning that the conclusions of the present study remain intact. Nonlinear modeling, provides more insight on how exam scheduling impacts exam grades. This insight may be used to design better exam schedules. "Better" in this case means that the schedule provides a fair chance for students to write their exams.

Some findings we concluded from the above spline plots:

- We can observe that PRE has a more complicated structure.
- Initially, increasing the amount of days a student has for studying improves the student's final exam grades (shown by the positive increase). However, after 5 days there is decline in performance.
- the plot of DAYS_FR_START shows a periodic pattern. This pattern suggests that students perform better at the beginning, middle, and end of the exam season.

Other findings and discussion:

- We also plotted the residuals of our model to test for normality (an assumption within mixed effect models). There is a slight departure from normality (right skew), however this only affects our type II error (false negative) and does not largely impact our conclusion.

CONCLUSION AND FUTURE WORK

Our analysis has shown that there is sufficient evidence that final exam scheduling has an impact on final exam grades. For example, our modeling has suggested that if a student were given one more day to study for a final exam, then their grade would move up 0.59%, after controlling numerous modes of variability. Limitations within our study suggest that a non-linear analysis with few normality assumptions should be performed next, to reasonably quantify the complete effect of exam scheduling. Having an accurate description of the effect of exam scheduling may assist with the creation of exam schedules, and may help redefine exam-hardship.

ACKNOWLEDGEMENTS

We'd like to give special thanks to Dr. Jackie Stewart and Dr. Gregory Dake, as this project would not have existed without their initiative.

We are grateful for Gabriela V. Cohen-Freue, our professor, for organizing STAT 450/550 and for her very informative lectures.

Special thanks to our STAT 550 team-mates: Joe Watson, Wayne Wang, and Atcher Zhang, for being top-notch mentors.

Lastly, we'd like to acknowledge our TA Dustin Johnson for his engaging R tutorials, and Estella Qi for her lectures on statistics communication.